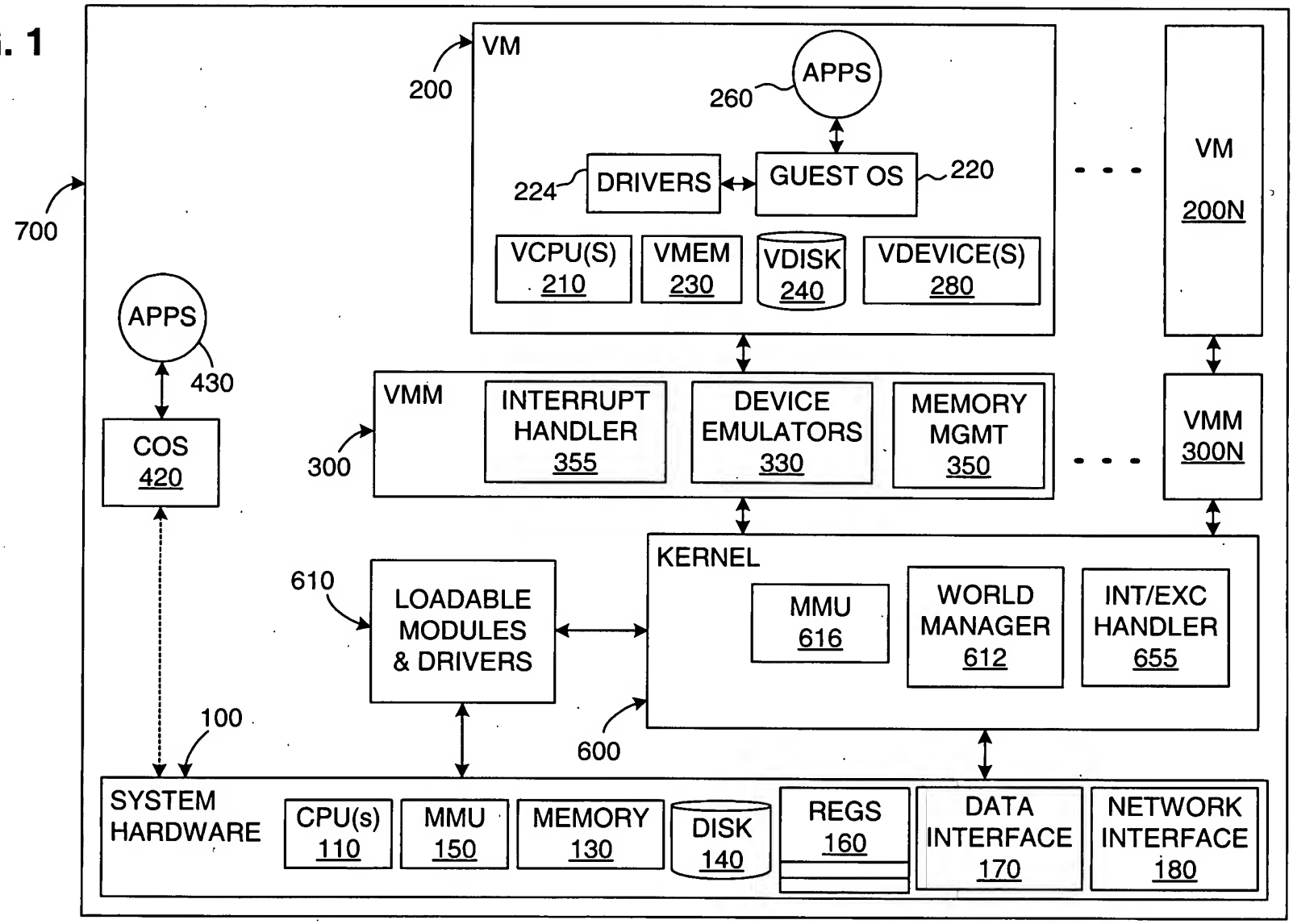
FIG. 1


700
VM
200
APPS
260
DRIVERS
224
GUEST OS
220
VCPU(S)
210
VMEM
230
VDISK
240
VDEVICE(S)
280
VM
200N
APPS
430
COS
420
VMM
300
INTERRUPT HANDLER
355
DEVICE EMULATORS
330
MEMORY MGMT
350
VMM
300N
KERNEL
610
LOADABLE MODULES & DRIVERS
MMU
616
WORLD MANAGER
612
INT/EXC HANDLER
655
100
600
SYSTEM HARDWARE
CPU(s)
110
MMU
150
MEMORY
130
DISK
140
REGS
160
DATA INTERFACE
170
NETWORK INTERFACE
180

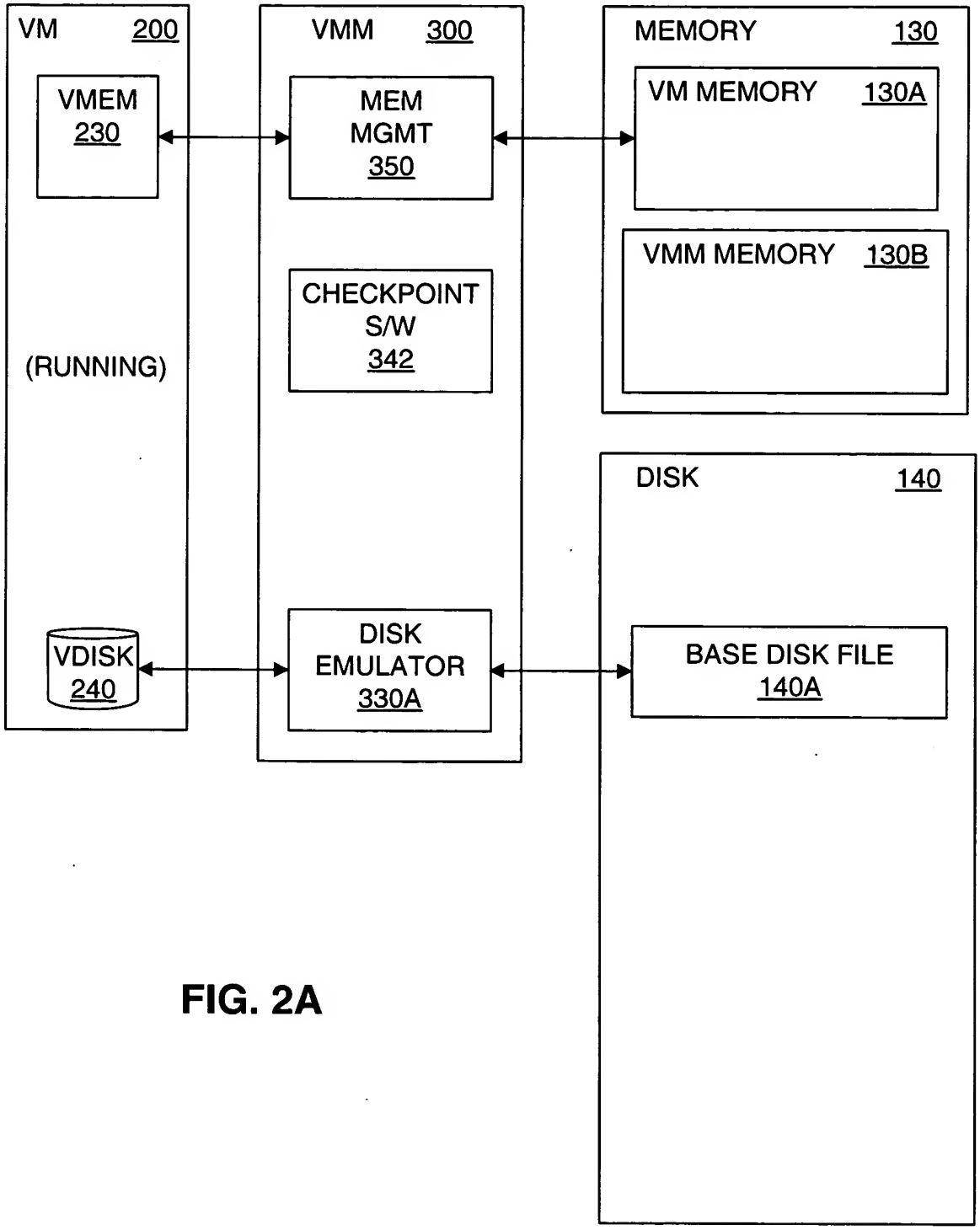

VM 200
VMEM 230
(RUNNING)
VDISK 240
VMM 300
MEM MGMT 350
CHECKPOINT S/W 342
DISK EMULATOR 330A
MEMORY 130
VM MEMORY 130A
VMM MEMORY 130B
DISK 140
BASE DISK FILE 140A

FIG. 2A

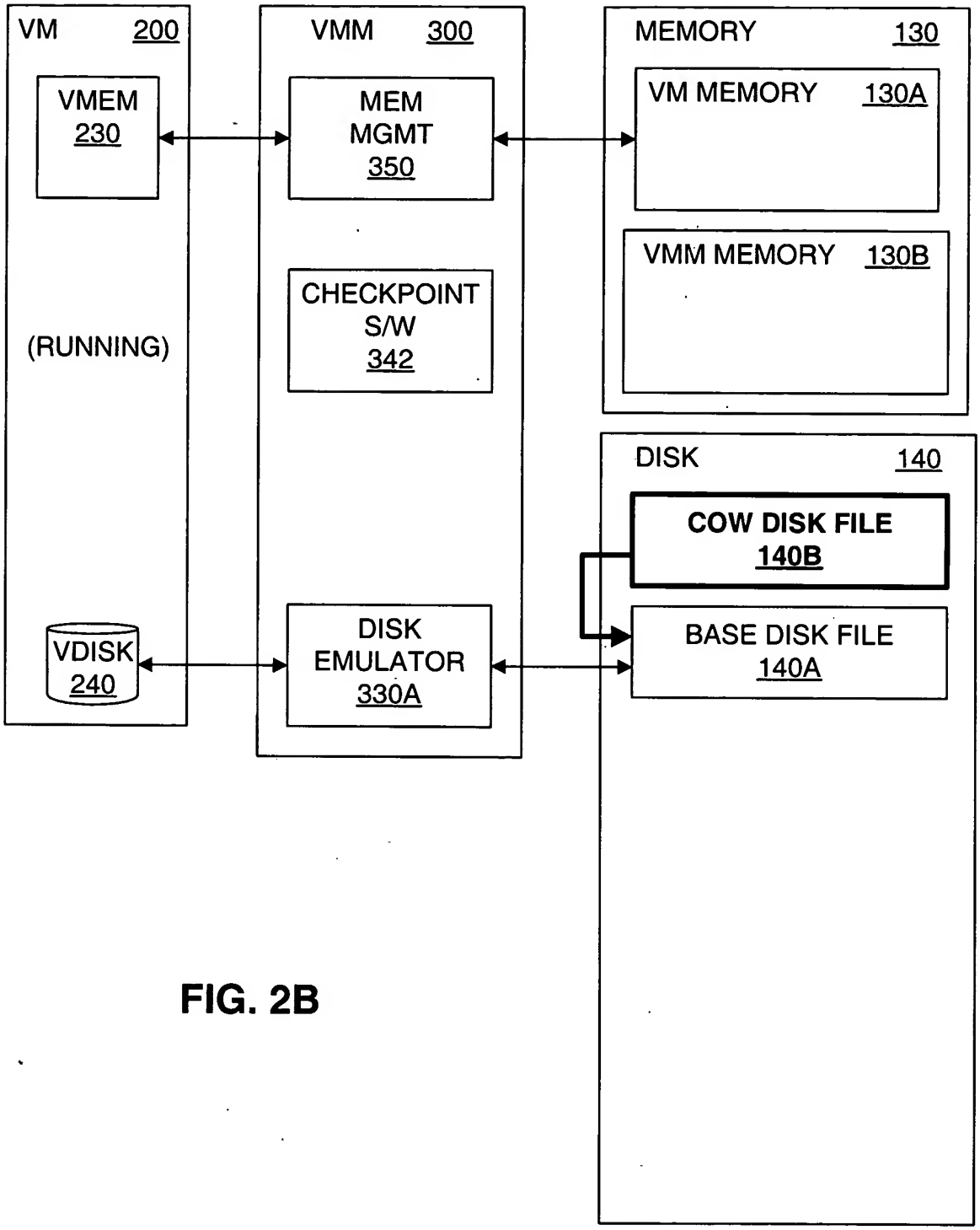

VM 200
VMEM 230
(RUNNING)
VDISK 240
VMM 300
MEM MGMT 350
CHECKPOINT S/W 342
DISK EMULATOR 330A
MEMORY 130
VM MEMORY 130A
VMM MEMORY 130B
DISK 140
COW DISK FILE 140B
BASE DISK FILE 140A

FIG. 2B

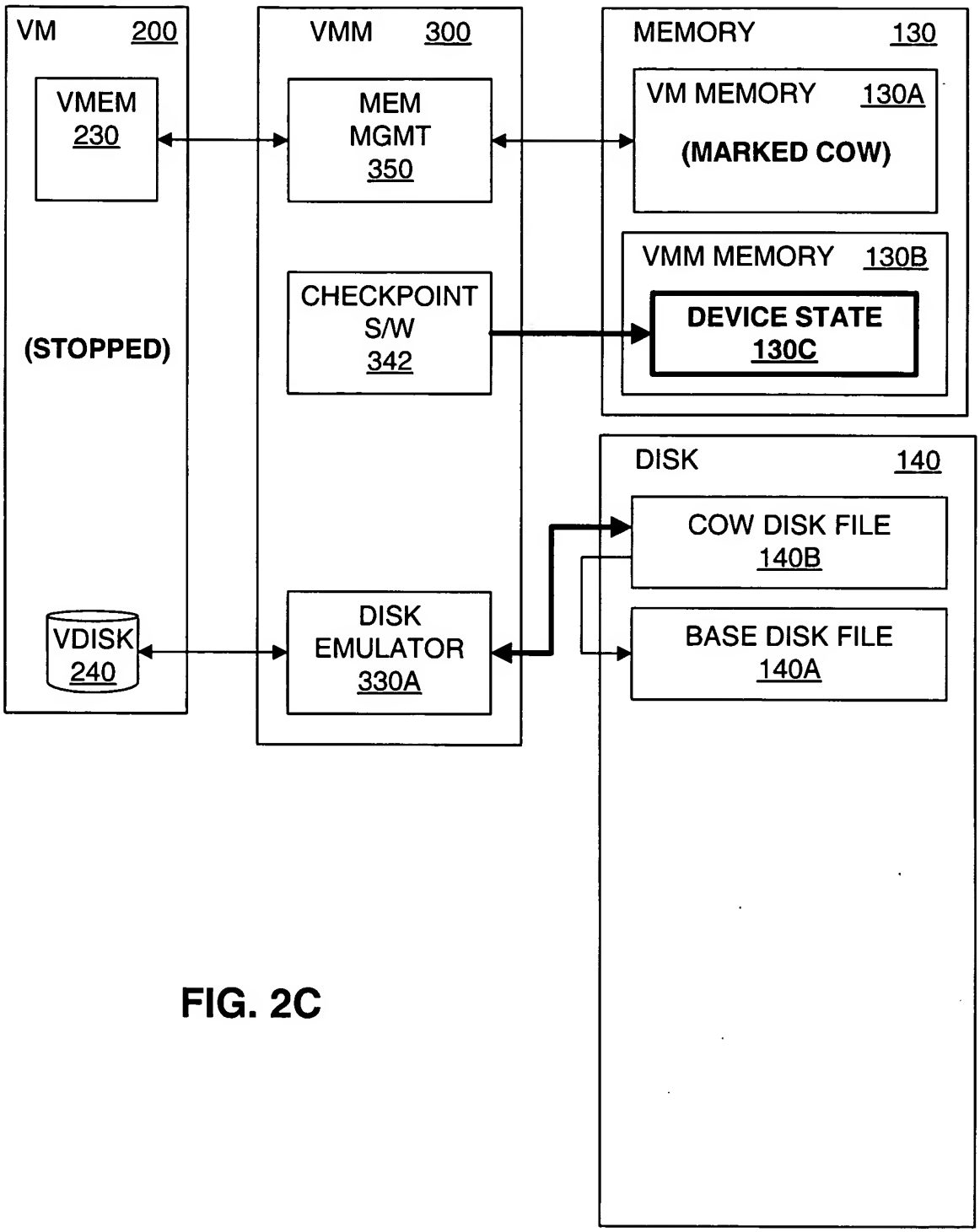

VM 200
VMEM 230
(STOPPED)
VDISK 240
VMM 300
MEM MGMT 350
CHECKPOINT S/W 342
DISK EMULATOR 330A
MEMORY 130
VM MEMORY 130A
(MARKED COW)
VMM MEMORY 130B
DEVICE STATE 130C
DISK 140
COW DISK FILE 140B
BASE DISK FILE 140A

FIG. 2C

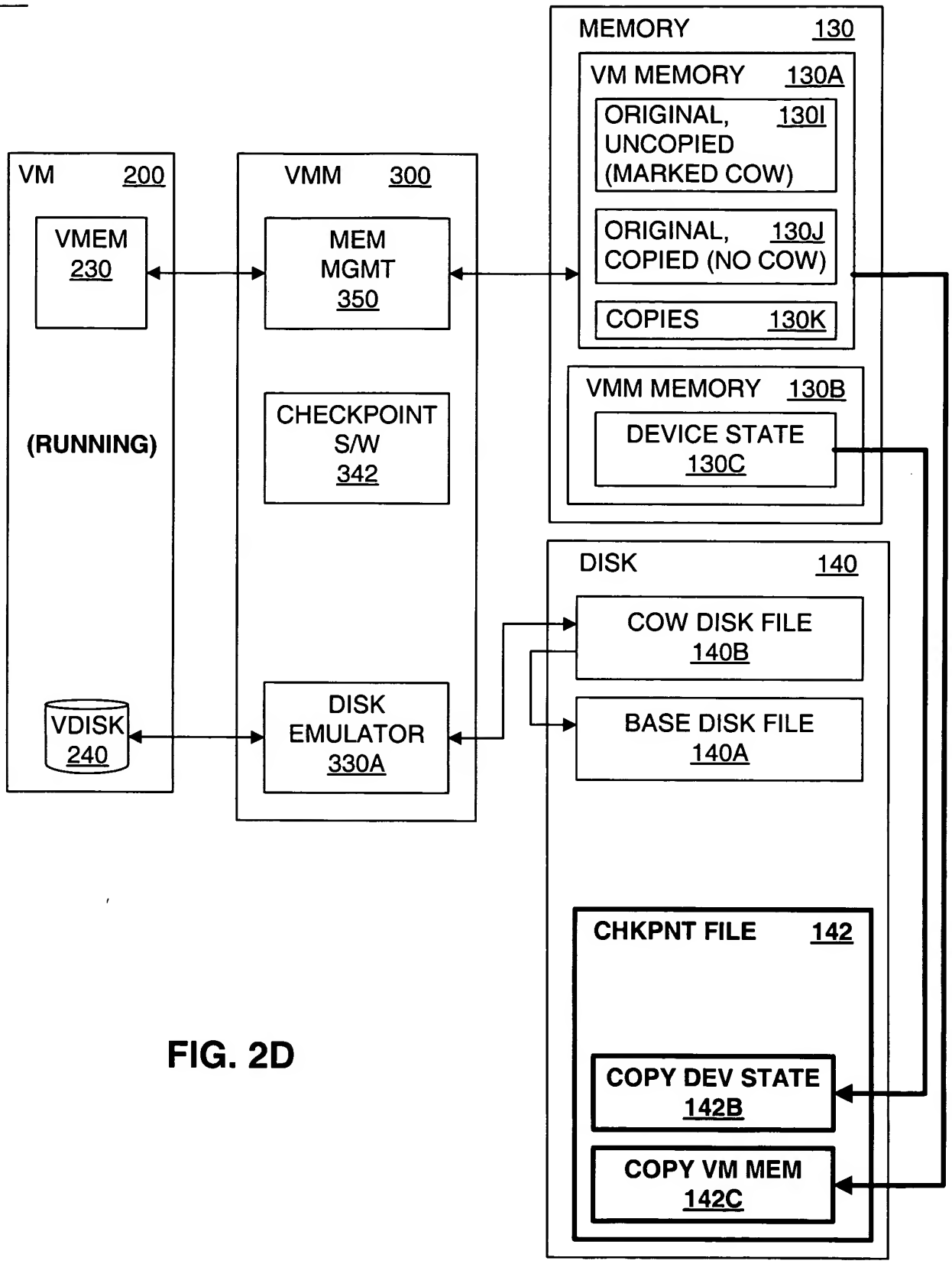

MEMORY 130
VM MEMORY 130A
ORIGINAL, UNCOPIED (MARKED COW) 130I
ORIGINAL, COPIED (NO COW) 130J
COPIES 130K
VMM MEMORY 130B
DEVICE STATE 130C
VM 200
VMEM 230
(RUNNING)
VDISK 240
VMM 300
MEM MGMT 350
CHECKPOINT S/W 342
DISK EMULATOR 330A
DISK 140
COW DISK FILE 140B
BASE DISK FILE 140A
CHKPNT FILE 142
COPY DEV STATE 142B
COPY VM MEM 142C

FIG. 2D

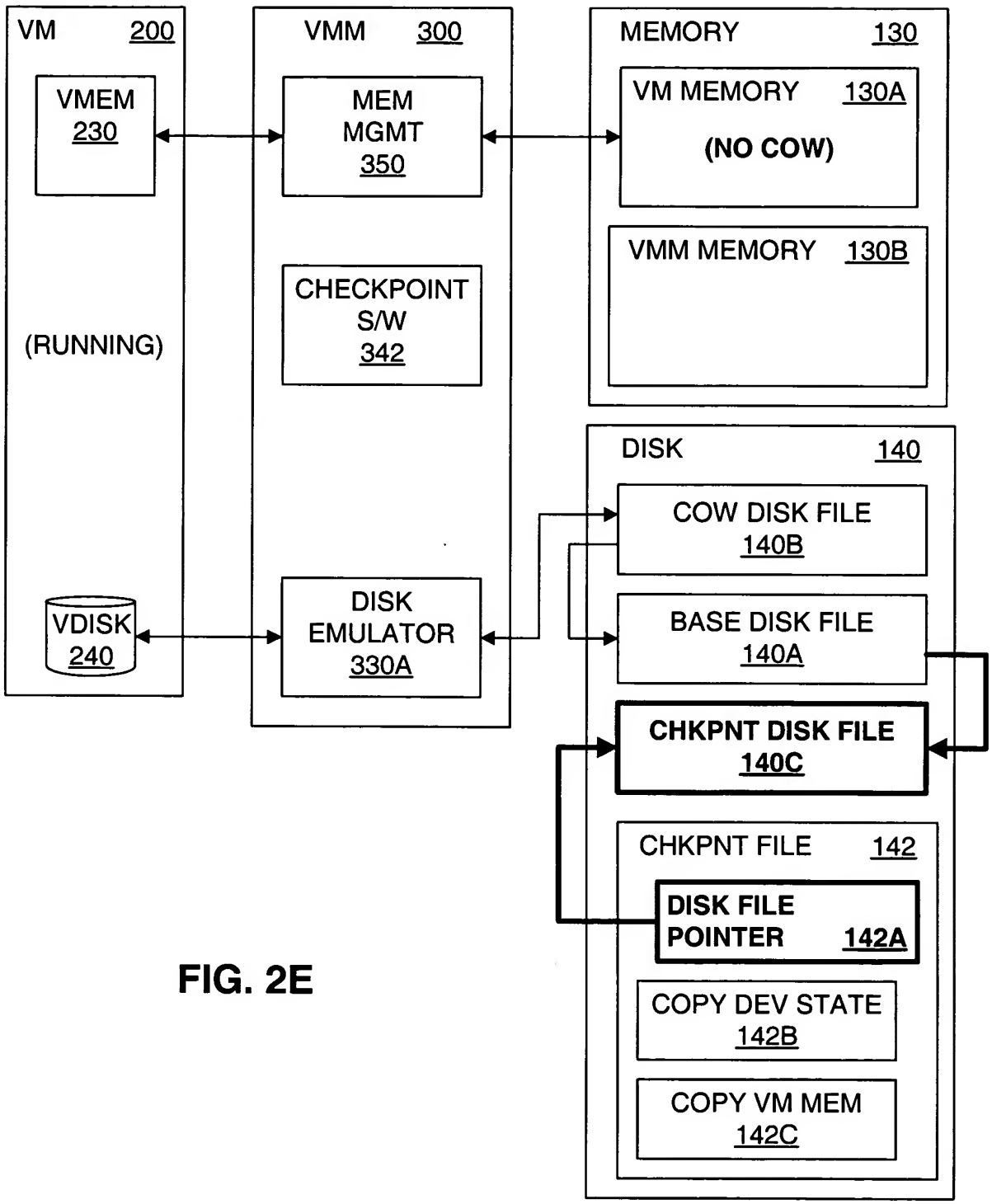

VM 200
VMEM 230
(RUNNING)
VDISK 240
VMM 300
MEM MGMT 350
CHECKPOINT S/W 342
DISK EMULATOR 330A
MEMORY 130
VM MEMORY 130A
(NO COW)
VMM MEMORY 130B
DISK 140
COW DISK FILE 140B
BASE DISK FILE 140A
CHKPNT DISK FILE 140C
CHKPNT FILE 142
DISK FILE POINTER 142A
COPY DEV STATE 142B
COPY VM MEM 142C

FIG. 2E

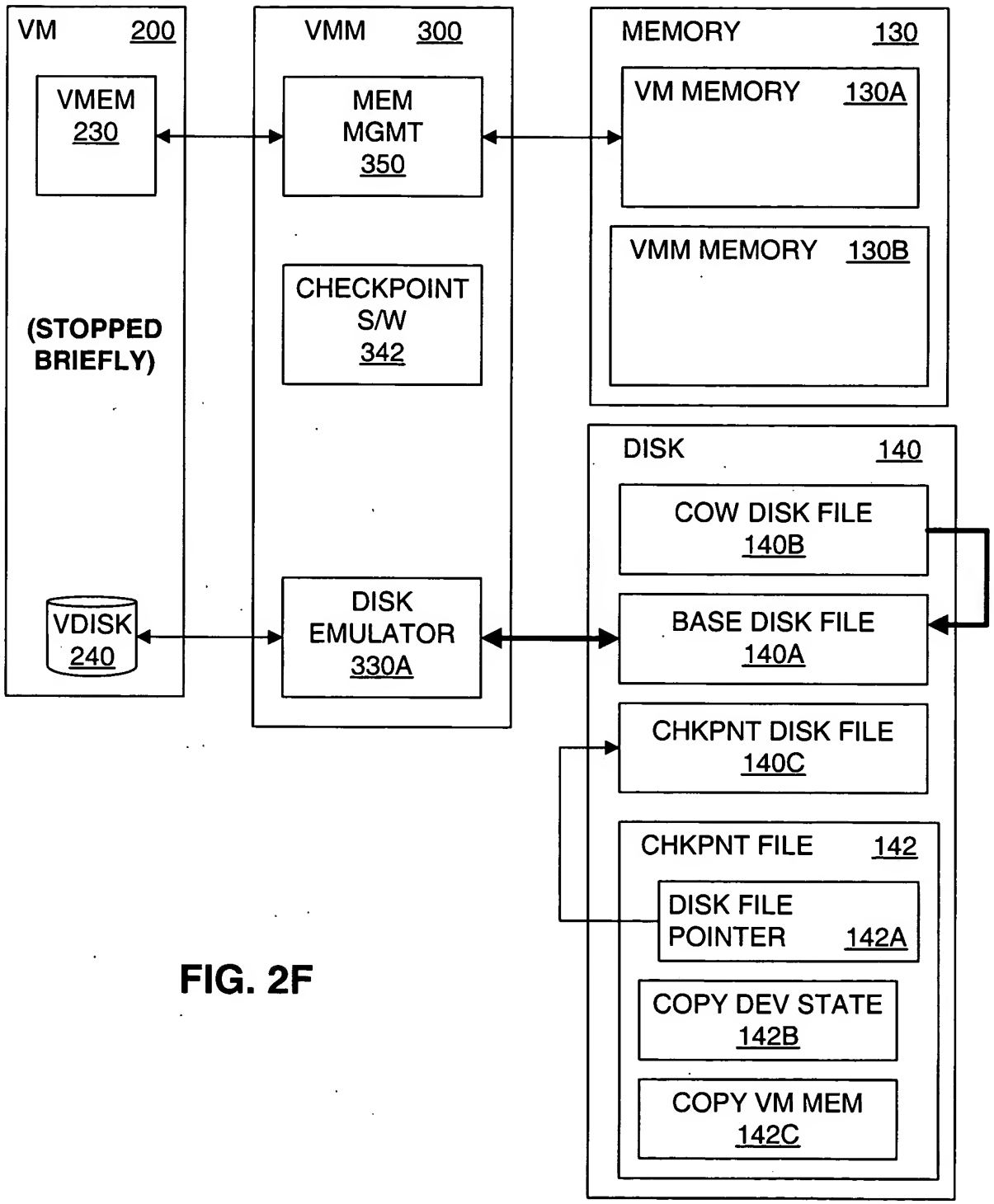

VM 200
VMEM 230
(STOPPED BRIEFLY)
VDISK 240
VMM 300
MEM MGMT 350
CHECKPOINT S/W 342
DISK EMULATOR 330A
MEMORY 130
VM MEMORY 130A
VMM MEMORY 130B
DISK 140
COW DISK FILE 140B
BASE DISK FILE 140A
CHKPNT DISK FILE 140C
CHKPNT FILE 142
DISK FILE POINTER 142A
COPY DEV STATE 142B
COPY VM MEM 142C

FIG. 2F

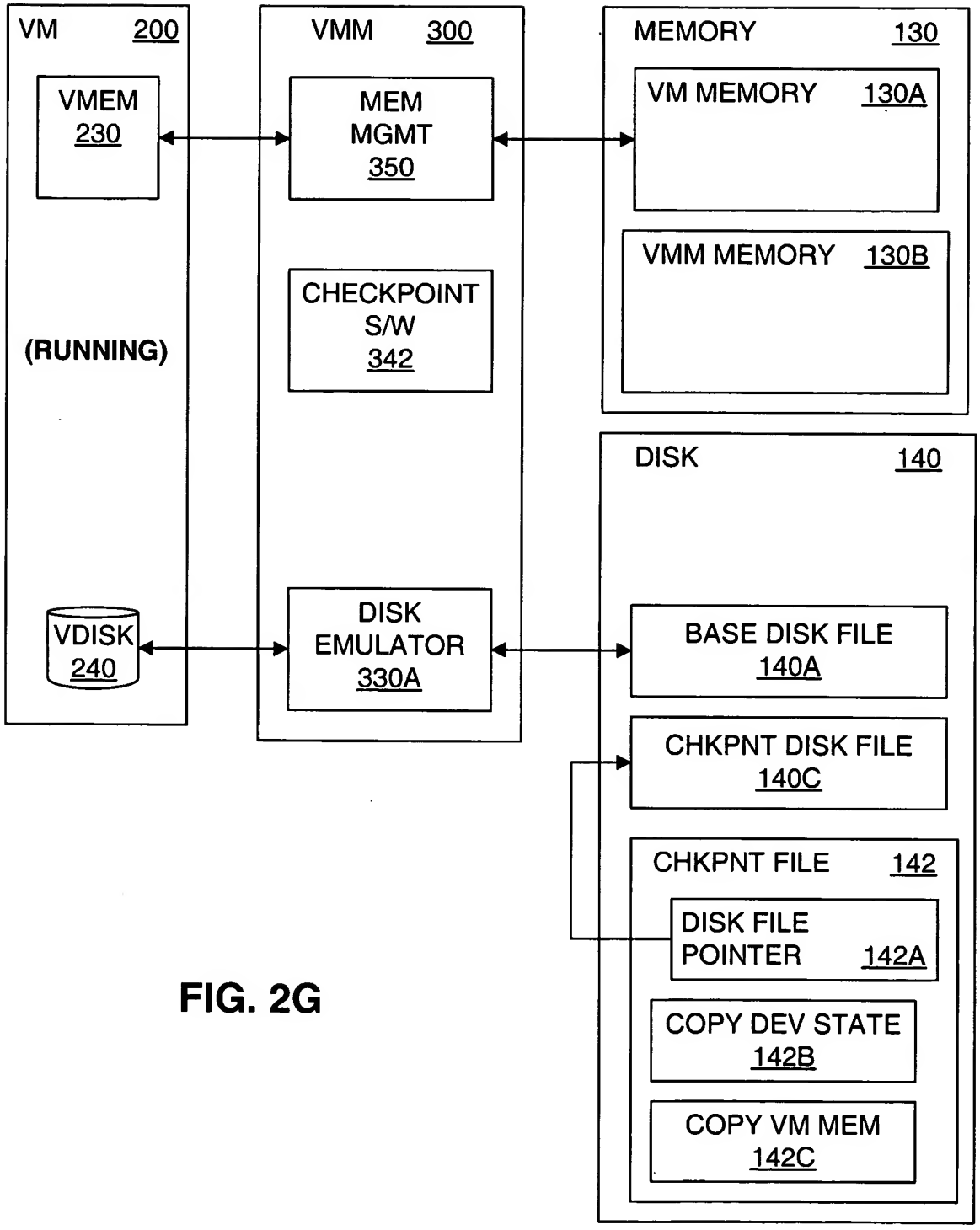

VM 200
VMEM 230
(RUNNING)
VDISK 240
VMM 300
MEM MGMT 350
CHECKPOINT S/W 342
DISK EMULATOR 330A
MEMORY 130
VM MEMORY 130A
VMM MEMORY 130B
DISK 140
BASE DISK FILE 140A
CHKPNT DISK FILE 140C
CHKPNT FILE 142
DISK FILE POINTER 142A
COPY DEV STATE 142B
COPY VM MEM 142C

FIG. 2G

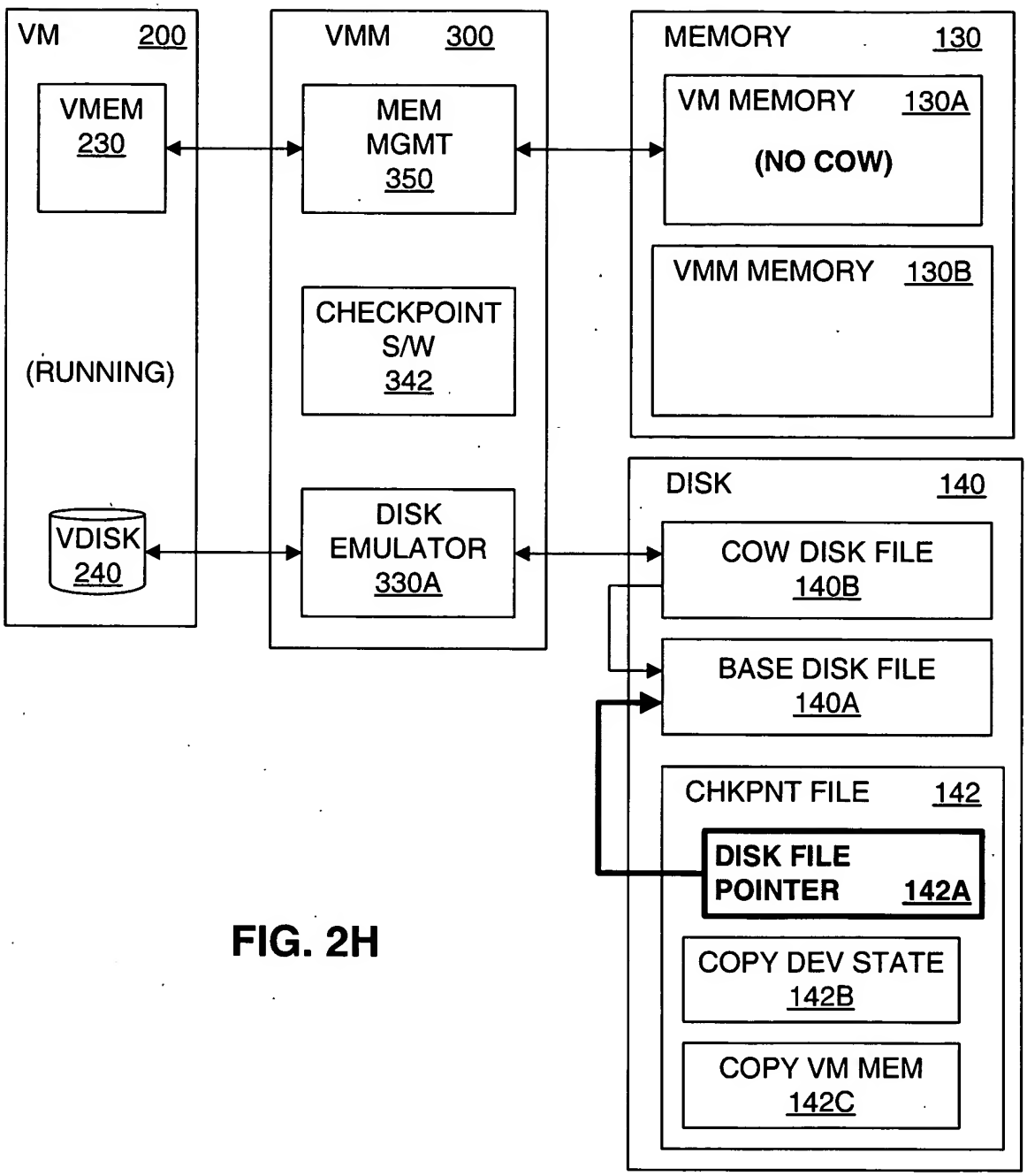

VM 200
VMEM 230
(RUNNING)
VDISK 240
VMM 300
MEM MGMT 350
CHECKPOINT S/W 342
DISK EMULATOR 330A
MEMORY 130
VM MEMORY 130A
(NO COW)
VMM MEMORY 130B
DISK 140
COW DISK FILE 140B
BASE DISK FILE 140A
CHKPNT FILE 142
DISK FILE POINTER 142A
COPY DEV STATE 142B
COPY VM MEM 142C

FIG. 2H

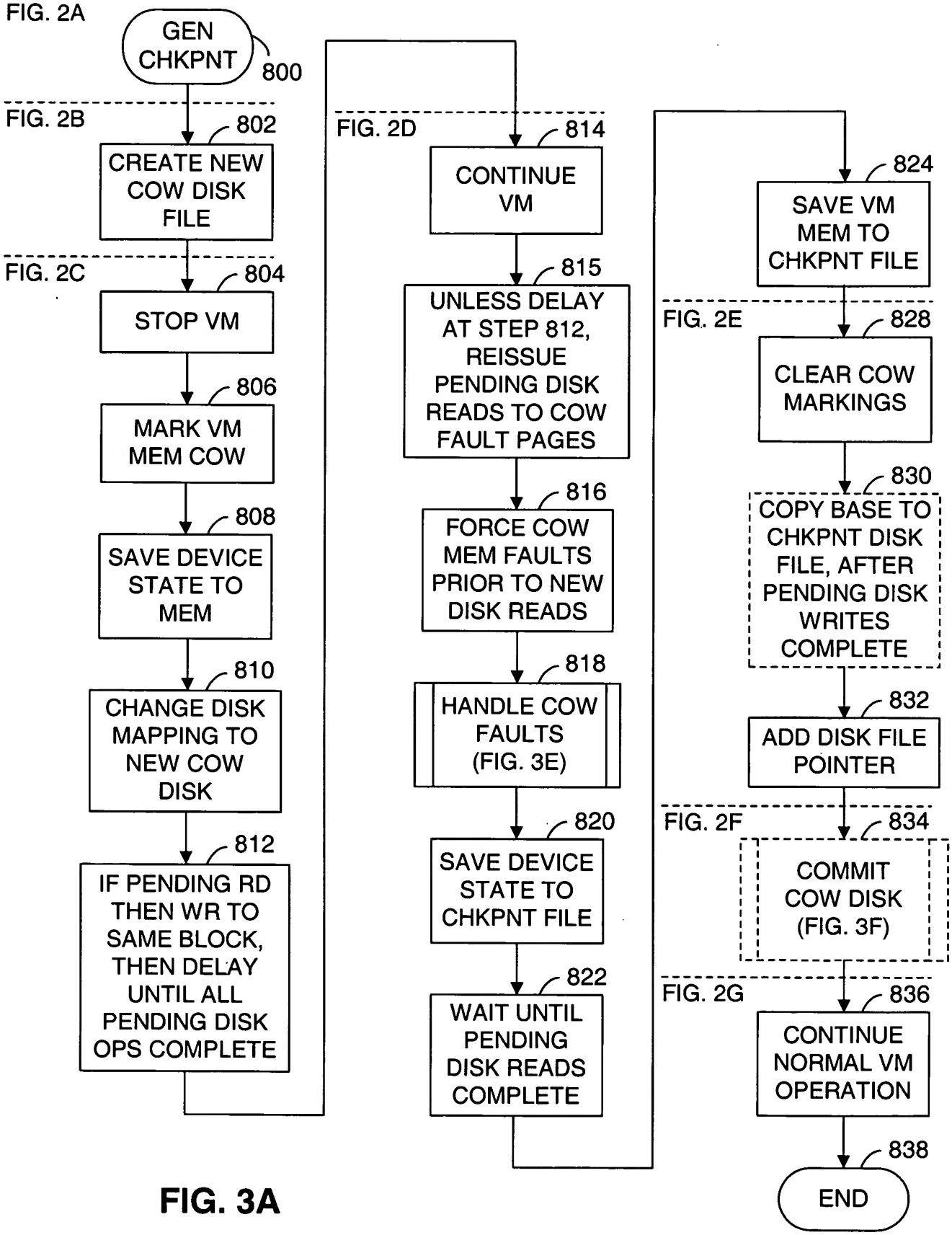


FIG. 2A
GEN CHKPNT
800
FIG. 2B
802
CREATE NEW COW DISK FILE
FIG. 2C
804
STOP VM
806
MARK VM MEM COW
808
SAVE DEVICE STATE TO MEM
810
CHANGE DISK MAPPING TO NEW COW DISK
812
IF PENDING RD THEN WR TO SAME BLOCK, THEN DELAY UNTIL ALL PENDING DISK OPS COMPLETE
FIG. 2D
814
CONTINUE VM
815
UNLESS DELAY AT STEP 812, REISSUE PENDING DISK READS TO COW FAULT PAGES
816
FORCE COW MEM FAULTS PRIOR TO NEW DISK READS
818
HANDLE COW FAULTS (FIG. 3E)
820
SAVE DEVICE STATE TO CHKPNT FILE
822
WAIT UNTIL PENDING DISK READS COMPLETE
824
SAVE VM MEM TO CHKPNT FILE
FIG. 2E
828
CLEAR COW MARKINGS
830
COPY BASE TO CHKPNT DISK FILE, AFTER PENDING DISK WRITES COMPLETE
832
ADD DISK FILE POINTER
FIG. 2F
834
COMMIT COW DISK (FIG. 3F)
FIG. 2G
836
CONTINUE NORMAL VM OPERATION
838
END

FIG. 3A

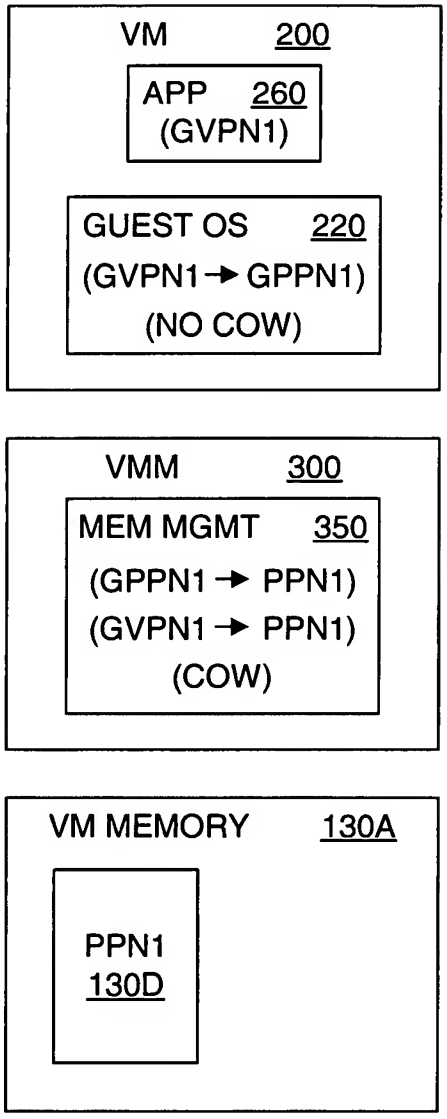

VM 200
APP 260
(GVPN1)
GUEST OS 220
(GVPN1 → GPPN1)
(NO COW)
VMM 300
MEM MGMT 350
(GPPN1 → PPN1)
(GVPN1 → PPN1)
(COW)
VM MEMORY 130A
PPN1
130D

FIG. 3B

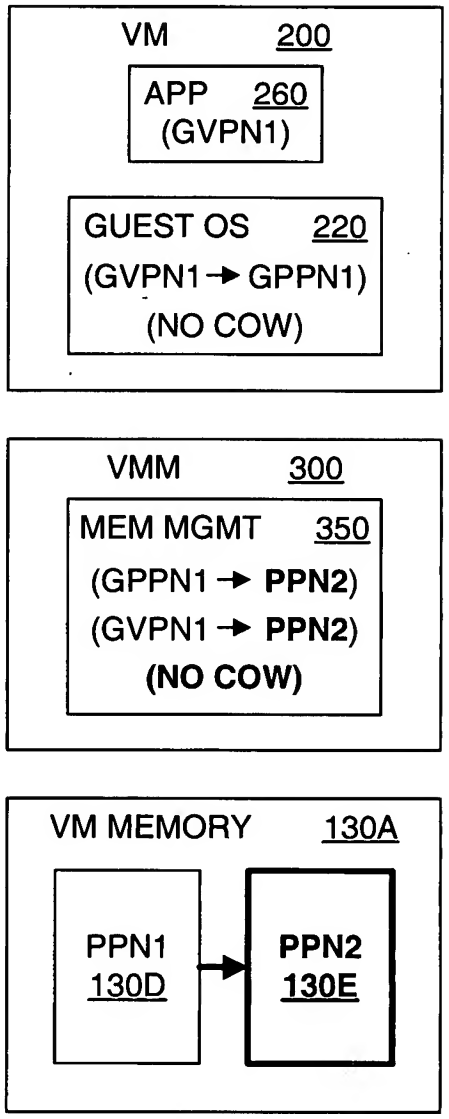

VM 200
APP 260
(GVPN1)
GUEST OS 220
(GVPN1 → GPPN1)
(NO COW)
VMM 300
MEM MGMT 350
(GPPN1 → PPN2)
(GVPN1 → PPN2)
(NO COW)
VM MEMORY 130A
PPN1
130D
PPN2
130E

FIG. 3C

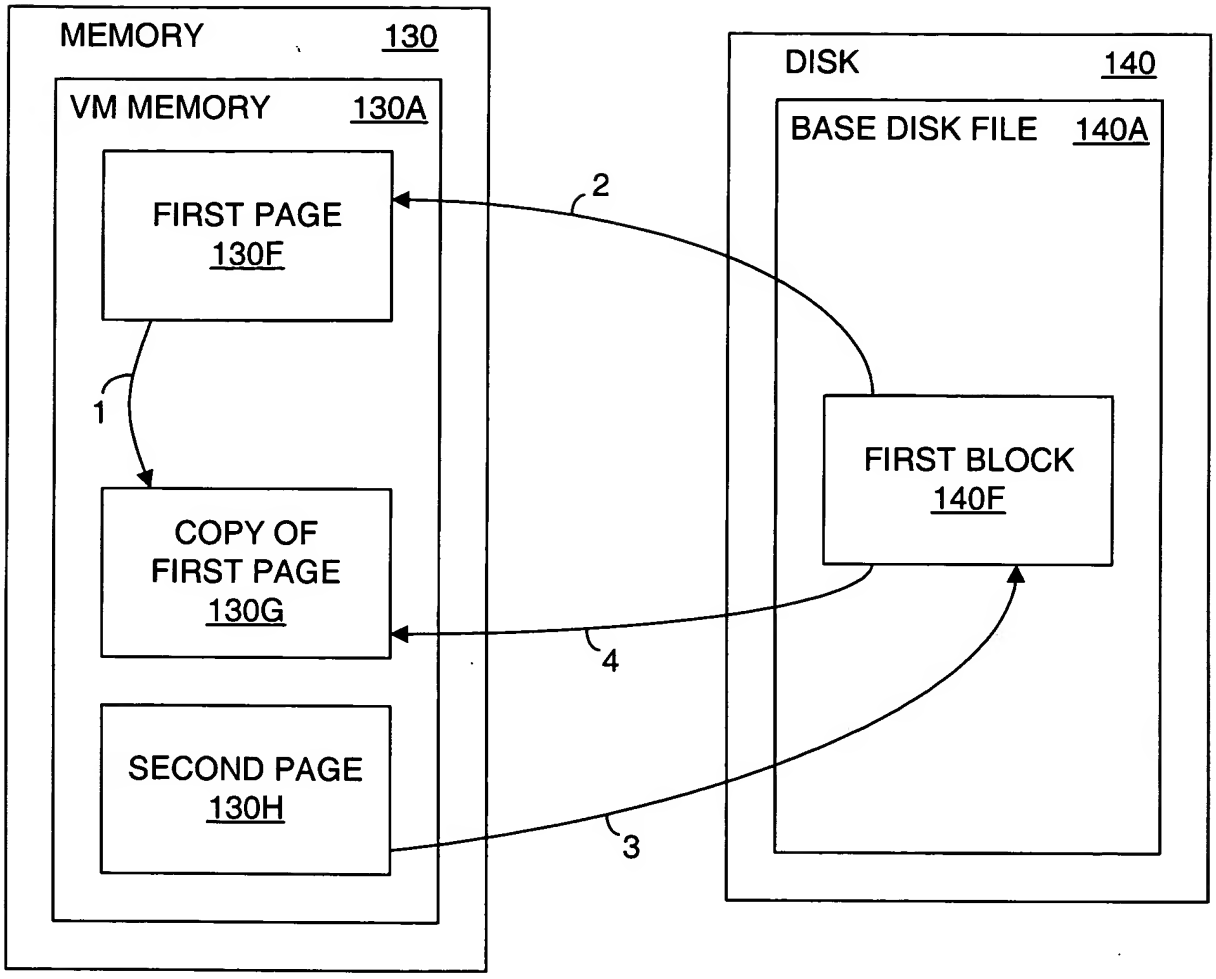

MEMORY
130
VM MEMORY
130A
FIRST PAGE
130F
COPY OF
FIRST PAGE
130G
SECOND PAGE
130H
DISK
140
BASE DISK FILE
140A
FIRST BLOCK
140F
1
2
3
4

FIG. 3D

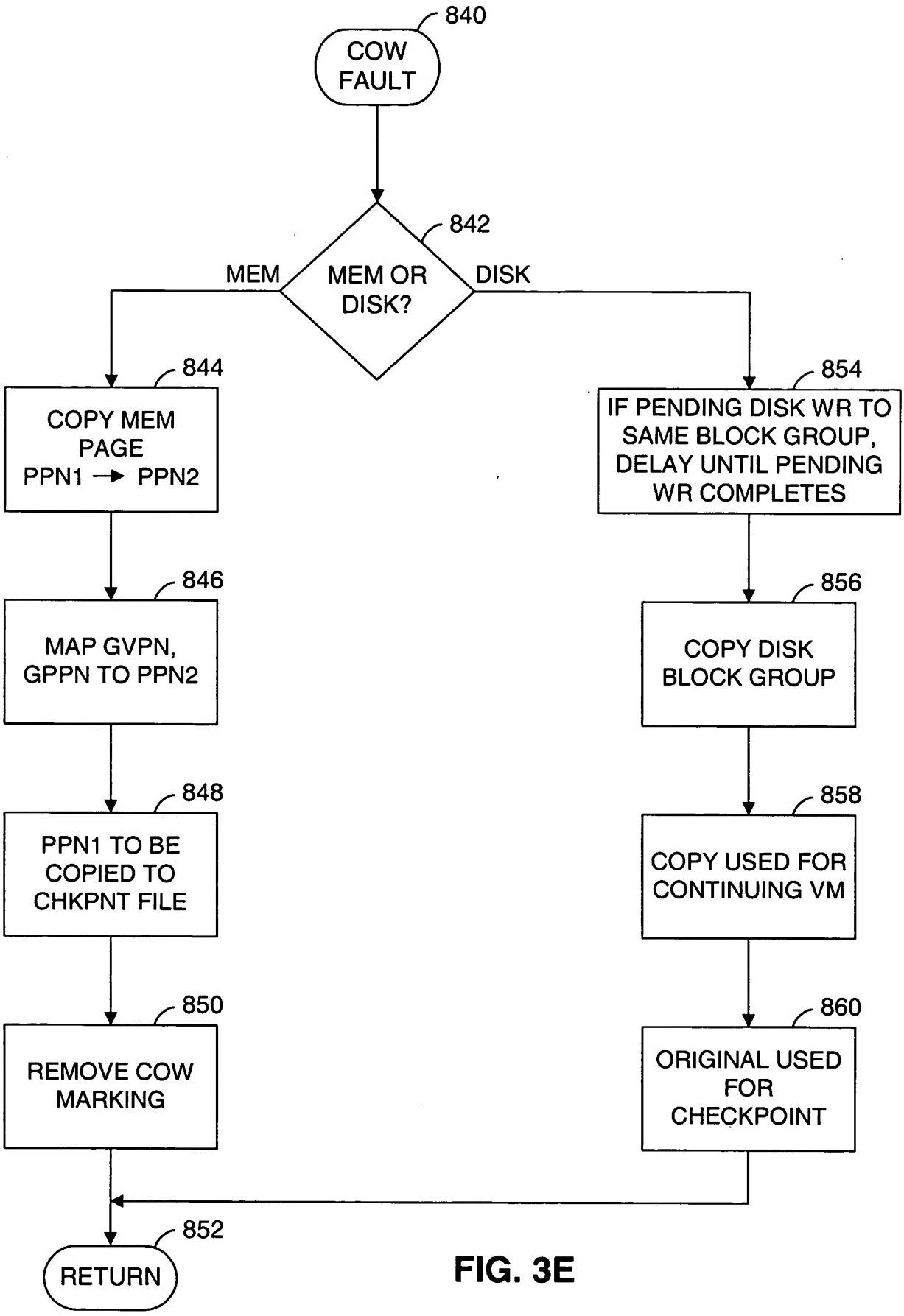

840
COW FAULT
842
MEM OR DISK?
MEM
DISK
844
COPY MEM PAGE PPN1 → PPN2
846
MAP GVPN, GPPN TO PPN2
848
PPN1 TO BE COPIED TO CHKPNT FILE
850
REMOVE COW MARKING
852
RETURN
854
IF PENDING DISK WR TO SAME BLOCK GROUP, DELAY UNTIL PENDING WR COMPLETES
856
COPY DISK BLOCK GROUP
858
COPY USED FOR CONTINUING VM
860
ORIGINAL USED FOR CHECKPOINT

FIG. 3E

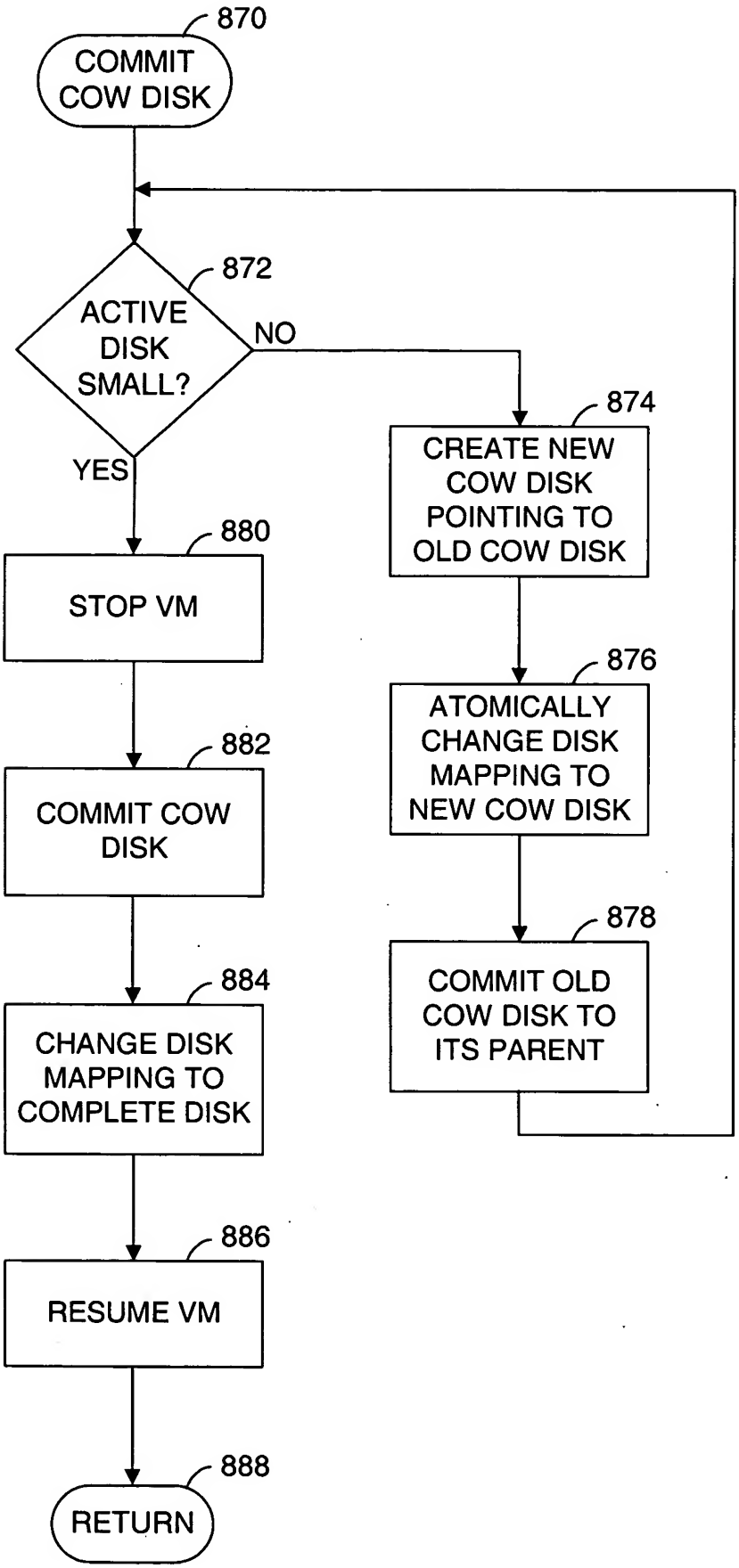

870
COMMIT COW DISK
872
ACTIVE DISK SMALL?
NO
YES
874
CREATE NEW COW DISK POINTING TO OLD COW DISK
876
ATOMICALLY CHANGE DISK MAPPING TO NEW COW DISK
878
COMMIT OLD COW DISK TO ITS PARENT
880
STOP VM
882
COMMIT COW DISK
884
CHANGE DISK MAPPING TO COMPLETE DISK
886
RESUME VM
888
RETURN

FIG. 3F

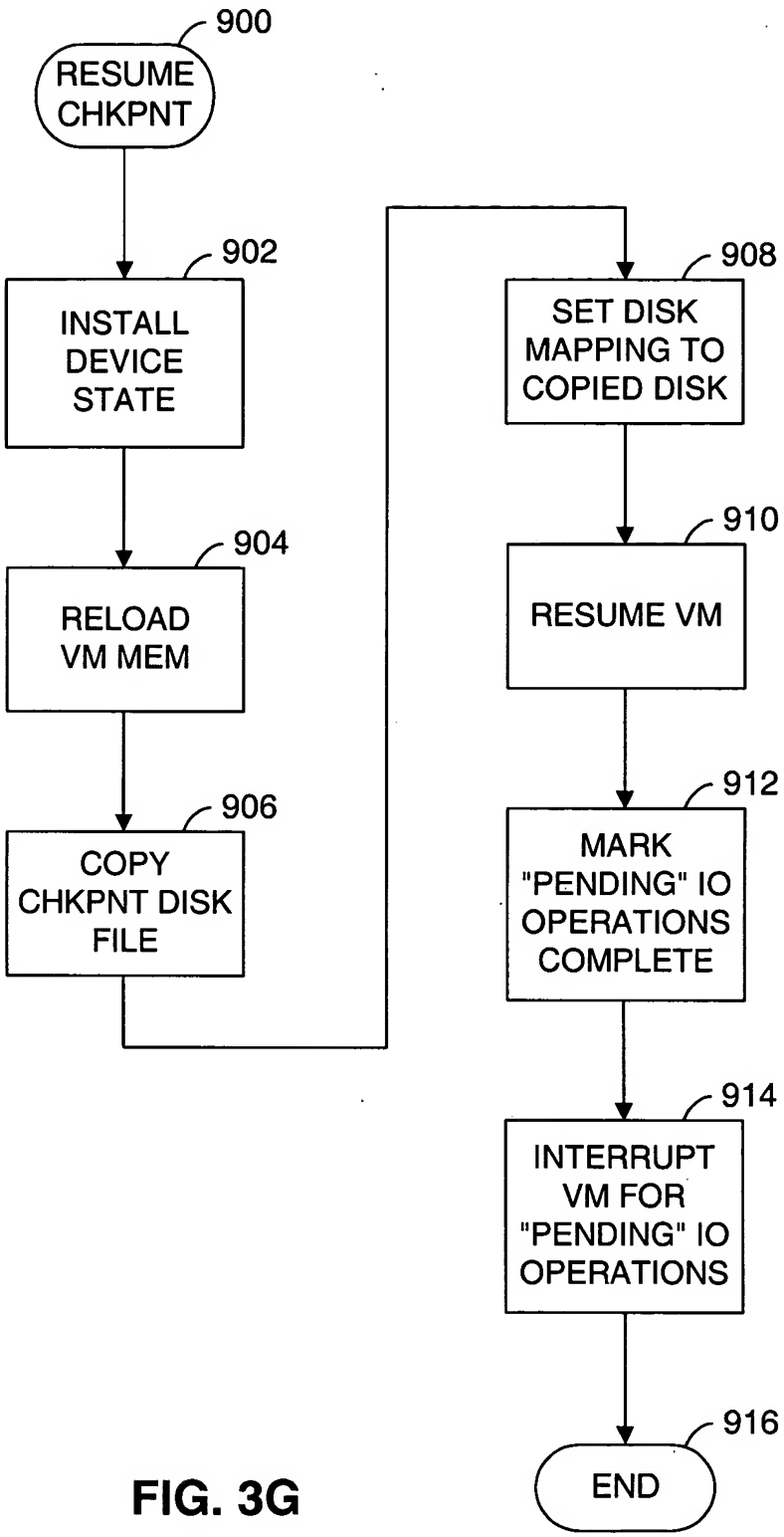

900
RESUME CHKPNT
902
INSTALL DEVICE STATE
904
RELOAD VM MEM
906
COPY CHKPNT DISK FILE
908
SET DISK MAPPING TO COPIED DISK
910
RESUME VM
912
MARK "PENDING" IO OPERATIONS COMPLETE
914
INTERRUPT VM FOR "PENDING" IO OPERATIONS
916
END

FIG. 3G

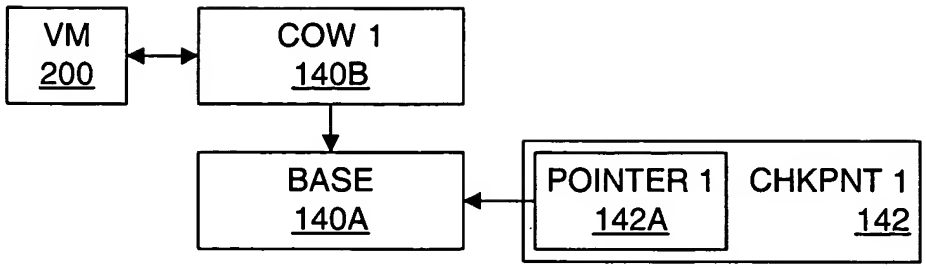


FIG. 4A

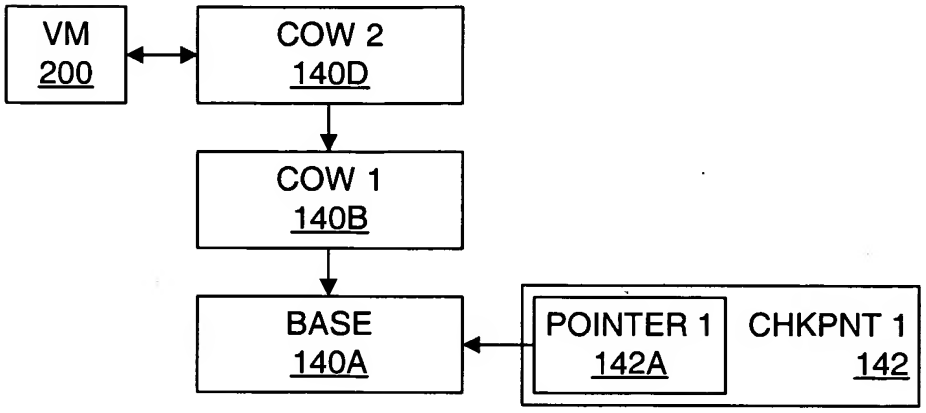


FIG. 4B

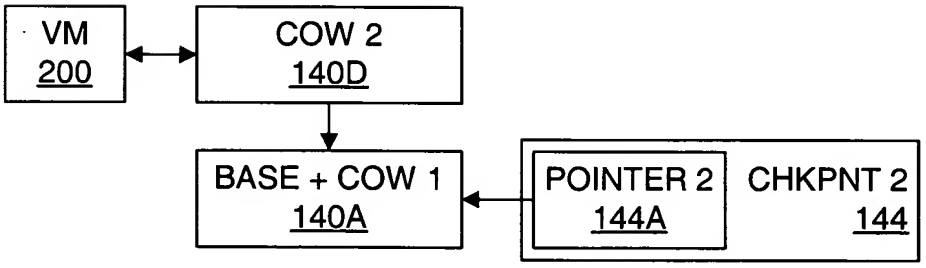


FIG. 4C

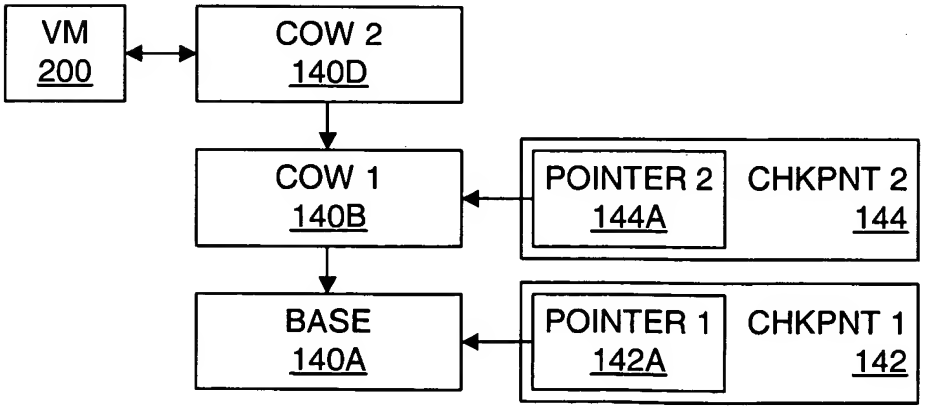


FIG. 4D

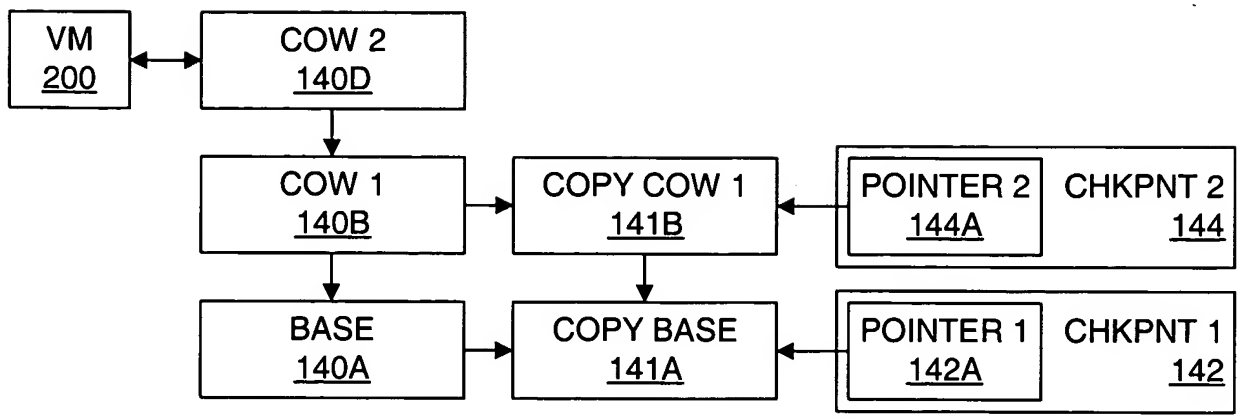


FIG. 5A

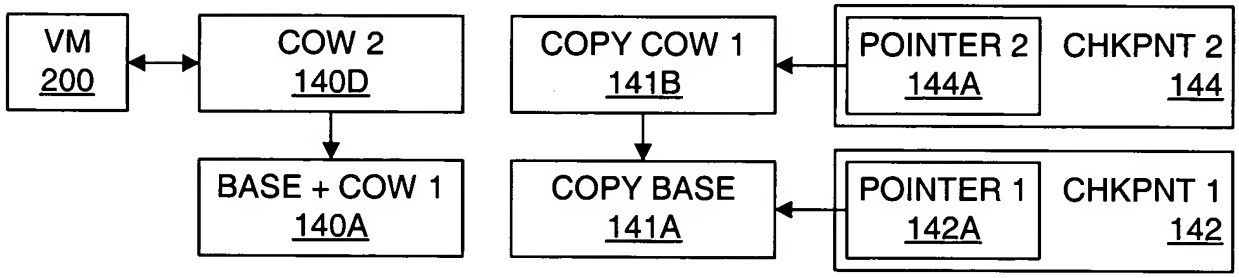


FIG. 5B

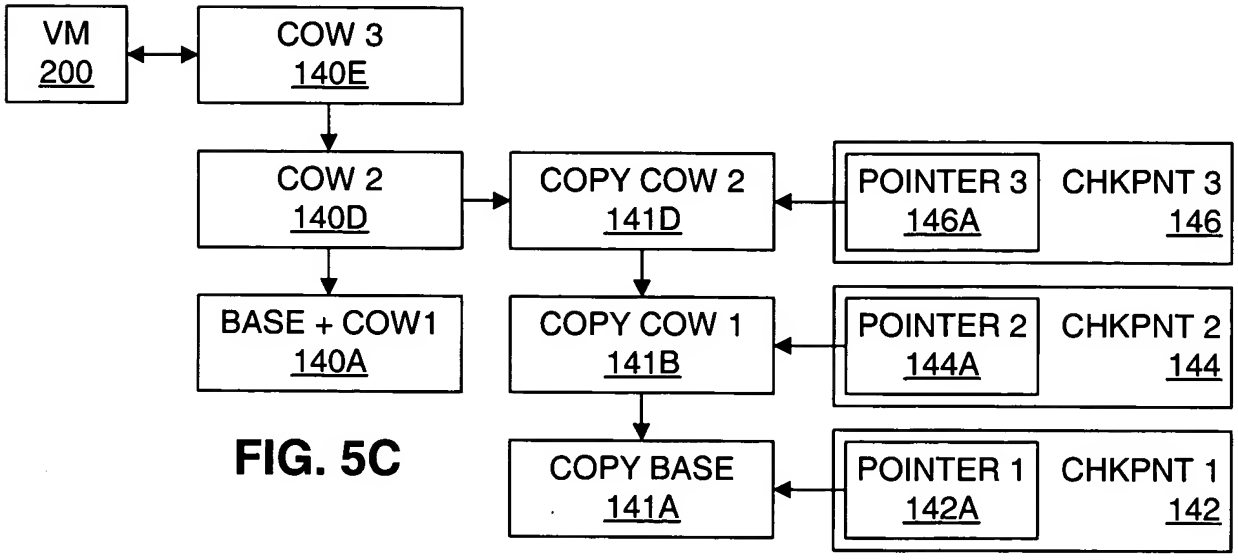


FIG. 5C

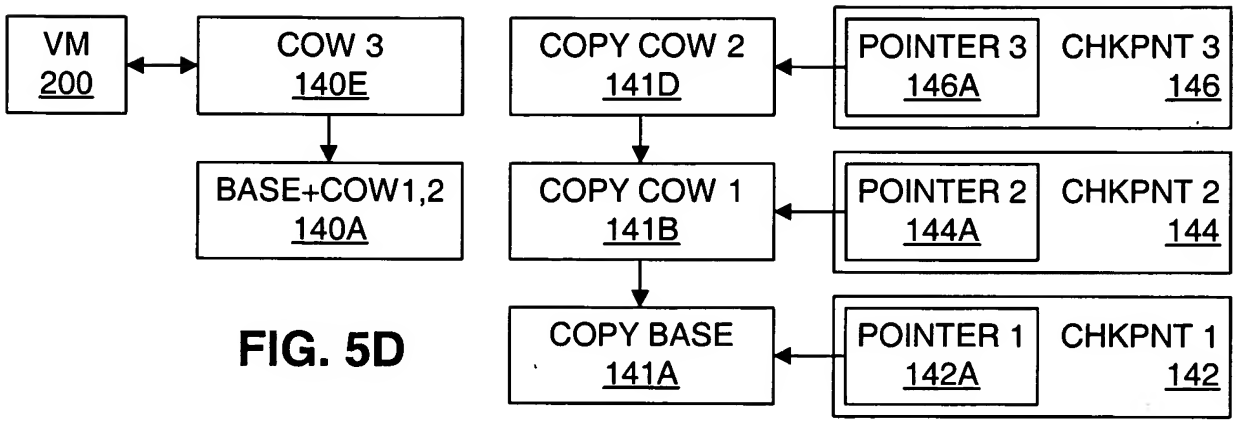

VM
200
COW 3
140E
BASE+COW1,2
140A
COPY COW 2
141D
COPY COW 1
141B
COPY BASE
141A
POINTER 3
146A
CHKPNT 3
146
POINTER 2
144A
CHKPNT 2
144
POINTER 1
142A
CHKPNT 1
142

FIG. 5D

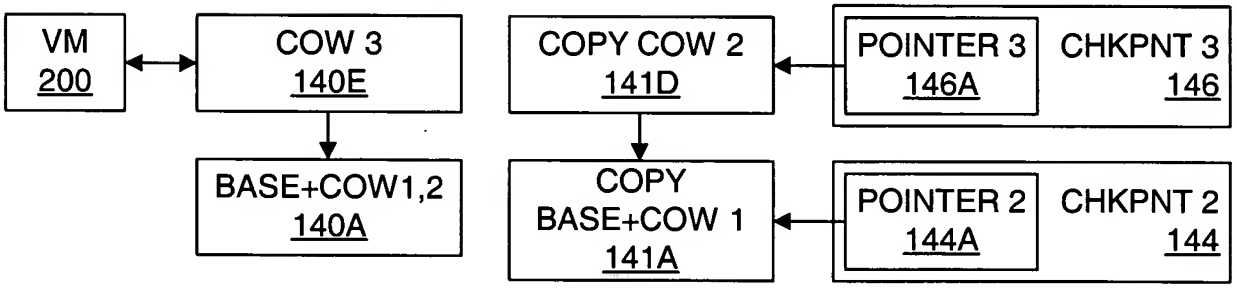

VM
200
COW 3
140E
BASE+COW1,2
140A
COPY COW 2
141D
COPY
BASE+COW 1
141A
POINTER 3
146A
CHKPNT 3
146
POINTER 2
144A
CHKPNT 2
144

FIG. 5E